A Comprehensive Epithelial Tubo-Ovarian Cancer Risk Prediction Model Incorporating Genetic and Epidemiological Risk Factors

Andrew Lee*[1], Xin Yang*[1], Jonathan Tyrer[2], Aleksandra Gentry-Maharaj[3], Andy Ryan[3], Nasim Mavaddat[1], Alex P. Cunningham[1], Tim Carver[1], Stephanie Archer[4], Goska Leslie[1], Jatinderpal Kalsi[5,6], Faiza Gaba[7], Ranjit Manchanda[7,8], Simon A. Gayther[9], Susan J. Ramus[10,11], Fiona M. Walter[4], Marc Tischkowitz[12], Ian Jacobs[5,10], Usha Menon[3], Douglas F. Easton[1,2], Paul P.D. Pharoah[1,2], Antonis C. Antoniou[†1]

[1] Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, The University of Cambridge, Strangeways Research Laboratory, Cambridge CB1 8RN, UK.

[2] Centre for Cancer Genetic Epidemiology, Department of Oncology, The University of Cambridge, Cambridge CB1 8RN, UK.

[3] MRC Clinical Trials Unit at UCL, Institute of Clinical Trials & Methodology, University College London, London, WC1V 6LJ, UK.

[4] The Primary Care Unit, Department of Public Health & Primary Care, The University of Cambridge, Cambridge, CB1 8RN, UK.

[5] Department of Women's Cancer, Institute for Women's Health, University College London, London, WC1E 6HU, UK.

[6] UCL Research Department of Epidemiology & Public Health, University College London, London, WC1E 7HB, UK.

[7] Wolfson Institute for Preventive Medicine, CRUK Barts Cancer Centre, Queen Mary University of London, London, EC1M 6BQ, UK.

[8] Department of Gynaecological Oncology, Barts Health NHS Trust, Royal London Hospital, London E1 1BB, UK.

[9] Center for Bioinformatics and Functional Genomics, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA.

[10] School of Women's and Children's Health, Faculty of Medicine, University of New South Wales, Sydney 2052, Australia.

[11] Adult Cancer Program, Lowy Cancer Research Centre, University of NSW, Sydney 2052, Australia.

[12] Department of Medical Genetics and National Institute for Health Research, Cambridge Biomedical Research Centre, The University of Cambridge, Cambridge, UK.

[*]joint first authors

[†]**Corresponding author**: Antonis C. Antoniou, Centre for Cancer Genetic Epidemiology, University of Cambridge, Strangeways Research Laboratory, Worts Causeway, Cambridge, CB1 8RN, UK, aca20@medschl.cam.ac.uk, +44 1223 748630

Word Count: 3915

Abstract

Background: Epithelial tubo-ovarian cancer (EOC) has high mortality partly due to late diagnosis. Prevention is available but may be associated with adverse effects. A multifactorial risk model based on known genetic and epidemiological risk factors (RFs) for EOC can help identify females at higher risk who could benefit from targeted screening and prevention.

Methods: We developed a multifactorial EOC risk model for females of European ancestry incorporating the effects of pathogenic variants (PVs) in *BRCA1, BRCA2*, *RAD51C*, *RAD51D* and *BRIP1*, a polygenic risk score (PRS) of arbitrary size, the effects of RFs and explicit family history (FH) using a synthetic model approach. The PRS, PV and RFs were assumed to act multiplicatively.

Results: Based on a currently available PRS for EOC that explains 5% of the EOC polygenic variance, the estimated lifetime risks under the multifactorial model in the general population vary from 0.5% to 4.6% for the 1st to 99th percentiles of the EOC risk-distribution. The corresponding range for females with an affected first-degree relative is 1.9% to 10.3%. Based on the combined risk distribution, 33% of *RAD51D* PV carriers are expected to have a lifetime EOC risk of less than 10%. RFs provided the widest distribution, followed by the PRS. In an independent partial model validation, absolute and relative 5-year risks were well-calibrated in quintiles of predicted risk.

Conclusion: This multifactorial risk model can facilitate stratification, in particular among females with FH of cancer and/or moderate- and high-risk PVs. The model is available via the CanRisk Tool (www.canrisk.org).

Keywords: ovarian cancer; risk prediction; rare-variants; polygenic risk score; model

INTRODUCTION

Epithelial tubo-ovarian cancer (EOC), the seventh most common cancer in females globally, is often diagnosed at late stage and is associated with high mortality. There were 7,443 new cases of EOC and 4,116 deaths from EOC annually in the UK in 2015-2017.[1] Early detection could lead to an early-stage diagnosis, enabling curative treatment and reducing mortality. Annual multimodal screening using a longitudinal serum CA125 algorithm in females from the general population resulted in significantly more females diagnosed with early-stage disease but without a significant reduction in mortality.[2] Four-monthly screening using the same multimodal approach also resulted in a stage shift in females at high risk (>10% lifetime risk of EOC).[3] Currently, risk-reducing bilateral salpingo-oophorectomy (RRSO), upon completion of their families, remains the most effective prevention option,[4] and it has been recently suggested that RRSO would be cost-effective in postmenopausal females at >4% lifetime EOC risk.[5, 6] Beyond surgical risk, bilateral oophorectomy may be associated with increased cardiovascular mortality[7] and a potential increased risk of other morbidities such as parkinsonism, dementia, cardiovascular disease and osteoporosis,[8, 9] particularly in those who do not take HRT.[10] Therefore, it is important to target such prevention approaches to those at increased risk, who are most likely to benefit.

Over the last decade, there have been significant advances in our understanding of susceptibility to EOC. After age, family history (FH) is the most important risk factor for the disease. Approximately 35% of the observed familial relative risk (FRR) can be explained by rare pathogenic variants (PVs) in the *BRCA1*, *BRCA2, RAD51C*, *RAD51D* and *BRIP1* genes.[11-14] Recent evidence suggests that *PALB2*, *MLH1*, *MSH2* and *MSH6* are also involved in the EOC genetic susceptibility.[14-16] Common variants, each of small effect, identified through

genome-wide association studies (GWAS),[17, 18] explain a further 4%. Several epidemiological risk factors (RFs) are also known to be associated with EOC risk, including use of menopausal hormone therapy (MHT), body mass index (BMI), history of endometriosis, use of oral contraception, tubal ligation and parity.[19-24] Despite these advances, those at high risk of developing EOC are currently identified mainly through FH of the disease or on the basis of having PVs in *BRCA1* and *BRCA2*. However, more specific risk prediction could be achieved by combining data on all known epidemiological and genetic risk factors. The published EOC prediction models consider either RFs[22, 23, 25] or common variants.[22, 26] No published EOC risk prediction model takes into account the simultaneous effects of the established EOC susceptibility genetic variants (rare and common), residual FH and other known RFs.

Using complex segregation analysis, we previously developed an EOC risk prediction algorithm that considered the effects of PVs in *BRCA1* and *BRCA2* and explicit FH of EOC and breast cancer (BC).[11] The algorithm modelled the residual, unexplained familial aggregation using a polygenic model that captured other unobserved genetic effects. The model did not explicitly include the effects of other established intermediate-risk PVs in genes such as *RAD51C*, *RAD51D* and *BRIP1,*[12-14, 27] which are now included on routine gene panel tests, the effects of recently developed EOC PRSs or the known RFs.

Here we present a methodological framework for extending this model to incorporate the explicit effects of PVs in *RAD51C*, *RAD51D* and *BRIP1* for which reliable age-specific EOC risk estimates are currently available, up-to-date polygenic risk scores (PRS) and the known EOC RFs (Table 1). We used this multifactorial model to evaluate the impact of negative predictive testing in families with rare PVs and to assess the extent of EOC risk stratification

that can be achieved in the general population, females with a FH of EOC and those carrying rare PVs. We evaluated the performance of a subset of this model in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS),[2] where females from the general population were followed up prospectively.

Table 1. Summary of components of the EOC risk model.

Risk factor group	Risk factor category	Comments
Family history	Explicit family history of ovarian and other cancers (breast, prostate, male breast, pancreatic).	Considers families of arbitrary size and structure, including affected and unaffected relatives.
Sex		Sex of all family members
Age		Ages at cancer diagnosis or current ages/age at death of family members.
Genetic Factors		
Rare truncating/pathogenic variants	*BRCA1*	
	BRCA2	
	RAD51D	
	RAD51C	
	BRIP1	
Common genetic variants	Polygenic risk score	Explaining 5.0% of the polygenic variance.
Unobserved genetic effects	Residual polygenic component	Accounts for the residual familial aggregation of EOC.
Lifestyle/hormonal/reproductive		
	Height	Measured in cm (5 categories)
	Body Mass Index	Measured in kg/m^2 (3 categories)
	Parity	Number of live births (3 categories)
	Endometriosis	Yes/No
	Use of oral contraception	Years of use (5 categories)
	Use of Hormone Replacement Therapy	Never/Ever
	Tubal Ligation	Yes/No
Breast tumour pathology	Oestrogen, Progesterone, HER2 Receptor, CK14, CK5/6 Status	As implemented in the BOADICEA breast cancer model
Demographic factors		
Country of Origin	Country	Defines the underlying incidences used.
Birth Cohort	Defined by the person's year of birth	8 calendar year-specific sets of incidences.
Family Ethnicity	Ashkenazi Jewish Origin	

METHODS

EOC risk prediction model development

No large datasets are currently available that include data on all known genetic and other EOC RFs. Therefore, we used a synthetic approach, described previously,[28] to extend our previous EOC model[11] by capitalising on published estimates of the associations of each RF with EOC. This approach was shown to provide valid risk estimates in the case of BC.[28-30]

Under the assumption that the effects of rare PVs, RFs and polygenic component are multiplicative on EOC risk, the incidence at age t for individual i was modelled as:

$$\lambda^{(i)}(t) = \lambda_0(t) \exp\left(\beta_{MG}^{(i)}(t) + \sum_{\rho} \boldsymbol{\beta}_{RF\rho}^{(i)}(t) \cdot \boldsymbol{z}_{RF\rho}^{(i)} + \beta_{PG}(t)\, x_P^{(i)}\right), \qquad (1)$$

$\lambda_0(t)$ is the baseline incidence. $\beta_{MG}^{(i)}(t)$ is the age-specific log-relative risk (log-RR) associated with individual i's PV carrier status (explained below), relative to the baseline. The log-RR for non-carriers is 0. $x_P^{(i)}$ is the polygenotype for individual i, assumed to follow a standard normal distribution in the general population and $\beta_{PG}(t)$ is the age-specific log-RR associated with the polygene, relative to the baseline incidence. $\boldsymbol{\beta}_{RF\rho}^{(i)}(t)$ is the array of the log-RRs associated with risk-factor ρ at age t, which may depend on PV carrier status, and $\boldsymbol{z}_{RF\rho}^{(i)}$ is the corresponding indicator array showing the category of risk-factor ρ for the individual. The baseline incidence was determined by constraining the overall incidences to agree with the population EOC incidence. To allow appropriately for missing RF information, only those RFs measured on a given individual are considered.

Major gene effects

To include the effects of *RAD51D, RAD51C* and *BRIP1*, we used the approach described previously where PVs in these genes were assumed to be risk alleles of a single MG locus.[31] A dominant model of inheritance was assumed for all rare PVs. To define the penetrance, we assumed the following order of dominance when an individual carried more than one PV (that is, the risk was determined by the highest risk PV and any lower risk PVs ignored): *BRCA1, BRCA2, RAD51D, RAD51C, BRIP1.*[31] The population allele frequencies for *RAD51D*, *RAD51C* and *BRIP1* and EOC RRs were obtained from published data (Table s3).[14, 27] Although PVs in *PALB2*, *MLH1*, *MSH2* and *MSH6* have been reported to be associated with EOC risk, PVs in *MLH1, MHS2* and *MSH6* are primarily associated with risk of specific subtypes of EOC (endometrioid and clear cell),[15] and at the time of development, precise EOC age-specific risk estimates for *PALB2* PV carriers were not available. Therefore, these were not considered at this stage.

Epidemiological Risk Factors

The RFs incorporated into the model include parity, use of oral contraception and MHT, endometriosis, tubal ligation, BMI and height. We assumed that the RFs were categorical and that an individual's category was fixed for their lifetime, although the RRs were allowed to vary with age. The RR estimates used in equation (1) and population distributions for each RF were obtained from large-scale external studies and from national surveillance data sources, using a synthetic approach as previously described.[28] Where possible, we used RR estimates that were adjusted for the other RFs included in the model and distributions from the UK. Details of the population distributions and RRs used in the model are given in Table s2. As in BOADICEA,[28] in order to decrease the runtime, we combined the RFs with

age-independent RRs into a single factor (specifically parity, tubal ligation, endometriosis, BMI and height).

Incorporating Polygenic Risk Scores

We included an EOC susceptibility PRS, assumed to form part of the polygene, using the methods previously developed.[11, 28] The polygenic component decomposes into a measured component due to the PRS (x_{PRS}) and an unmeasured component representing other familial effects (x_R):

$$x_P = x_{PRS} + x_R. \tag{2}$$

x_{PRS} summarises the effects of multiple common variants and is assumed normally distributed with mean 0 and variance α^2 in the general population, with $0 \leq \alpha \leq 1$. The parameter α^2 is the proportion of the overall polygenic variance (after excluding the effects of all MGs) explained by the PRS. x_R is normally distributed with mean 0 and variance $1 - \alpha^2$. The approach used to calculate α^2 is described in the Supplementary Material. This implementation allows the effect size of the PRS to be dynamically varied, allowing an arbitrary PRS.

Here, for illustrating the model's risk-stratification potential, we considered the latest validated EOC PRS developed by the Ovarian Cancer Association Consortium,[32] which is composed of 36 variants (Table s1) and has a log-variance of 0.099, accounting for 5.0% of the overall polygenic variance in the model. This 36 variant PRS was found to perform equally well as those comprised of more variants based on penalised regression or Bayesian approaches.[32]

Other model components

The previous version[11] modelled the incidence of EOC and first female BC. To align with BOADICEA,[28] the model was extended to take account of female contralateral BC and the associations of *BRCA1/2* PVs with pancreatic cancer, male BC and prostate (Supplementary Methods).

Model Validation

Study subjects

A partial model validation was carried out in a nested case-control sample of females of self-reported European ancestry participating in UKCTOCS. Based on the data available, we were able to validate the model on the basis of FH, PRS and RFs. Details of the UKCTOCS study design, blood sampling process, DNA extraction and processing, variant selection, genotyping and data processing are described in the Supplementary Methods and published elsewhere.[33] The study was approved by local ethical review committees. In summary, the following self-reported information was collected at recruitment and used for model validation: parity, use of oral contraception and MHT, tubal ligation, BMI and height (Table s4). As the study participants were genotyped for only 15 SNPs known at the time to be associated with EOC risk, it was not possible to use the more recently developed PRS for model validation. Instead, as the model can accommodate an arbitrary PRS, a PRS based on the on the 15 available SNPs was used[33] (Table s5), for which $\alpha^2 = 0.037$. The UKCTOCSS study participants were independent of the sets used to generate this PRS.[33] Study participants were not screened for PVs in *BRCA1, BRCA2, RAD51C, RAD51D* or *BRIP1.*

Pedigree construction

The UKCTOCS recruitment questionnaire collected only summary data on FH of BC and EOC. Since the risk algorithm uses explicit FH information, these data were used to reconstruct the pedigrees, which included information on incidences in the first- and second-degree relatives (Supplementary Methods).

Statistical analysis

All UKCTOCS participants were followed using electronic health record linkage to national cancer and death registries. For this study, they were censored at either: their age at EOC, their age at other (non-EOC) first cancer diagnosis, their age at death or age 79. To assess the model performance, a weighted approach was used whereby each participant was assigned a sampling weight based on the inverse of the probability of being included in the nested case-control study, given their disease status. Since all incident cancer cases were included, cases were assigned a weight of 1. The cases were matched to two random controls (females with no EOC cancer) recruited at the same regional centre, age at randomisation and year at recruitment.

We assessed the model calibration and discrimination of the predicted five-year risks. Females older than 74 years at entry were excluded. Cases that developed EOC beyond five years were treated as unaffected. For controls with a less than five-year follow-up, we predicted the EOC risks to the age at censoring. For all other controls and cases, we predicted five-year risks.

To assess model calibration, we partitioned the weighted sample into quintiles of predicted risk. Within each quintile, we compared the weighted mean of predicted risk to the weighted observed incidence using the Hosmer-Lemeshow (HL) chi-squared test.[34] To

assess RR calibration, the predicted and observed RRs were calculated relative to the corresponding means of risks over all quintiles. We also compared the expected (E) with the observed (O) EOC risk within the prediction interval by calculating the ratio of expected to observed cases (E/O). The 95% confidence interval (CI) for the ratio was calculated assuming a Poisson distribution.[35]

We assessed the model discrimination between females who developed and did not develop EOC within five years using the area under the ROC curve (AUC) (Supplementary Methods).

RESULTS

Model description

RAD51D, RAD51C and *BRIP1,* based on the assumed allele frequencies and RRs, account for 2.5% of the overall model polygenic variance. Figure 1 shows the predicted EOC risks for carriers of PVs in *BRCA1*, *BRCA2*, *RAD51D*, *RAD51C* and *BRIP1* for various FH scenarios. With unknown FH, the risks for carriers of PVs in *RAD51D, RAD51C* and *BRIP1* are 13%, 11% and 6%, respectively. For example, for a *BRIP1* PV carrier, the risk varies from 6% for a female without EOC FH to 18% for a female with two affected first-degree relatives. The model can also be used to predict risks in families in which PVs are identified but where other family members test negative (Figure s1). For females with a FH of EOC, the reduction in EOC risk after negative predictive testing is greatest for *BRCA1* PVs, with the risks being close to (though still somewhat greater than) population risk. This effect was most noticeable for females with a strong FH. Although a risk reduction is also seen for females whose mother carried a PV in *BRCA2*, *RAD51D*, *RAD51C* or *BRIP1*, the reduction is less marked. As expected, the predicted risks are still elevated compared to population.

Figures 2 and s2 show distributions of lifetime risk and risk by age 50 respectively for females untested for PVs, based on RFs and PRS, for two FH scenarios: (1) unknown FH (i.e. equivalent to a female from the general population); and (2) having a mother diagnosed with EOC at age 50. Table 2 shows the corresponding proportion of females falling into different risk categories. The variation in risk is greatest when including both the RFs and PRS. When considered separately, the distribution is widest for the RFs. Using the RFs and PRS combined, predicted lifetime risks vary from 0.5% for the 1st percentile to 4.6% for the 99th for a female with unknown FH and from 1.9% to 10.3% for a female with an affected mother.

Table 2. Percentage of females falling in different risk categories by status of PV in one of the high- or intermediate-risk genes included in the model and family history of cancer. The "population" risk category is defined as lifetime risk of <5% and risk to age 50 of <3%. "Moderate" risk category is defined as a lifetime risk of 5% or greater but less than 10% and a risk to age 50 of 3% or greater but less than 5%. "High" risk category is defined as a lifetime risk of 10% or greater and a risk to age 50 of 5% or greater. RF: epidemiological risk factors; PRS: polygenic risk score; NA: unknown family history; M 50: mother diagnosed with EOC at age 50. The population lifetime risk is 1.8%, and the population risk to age 50 is 0.27%.

PV status	Family history	Risk Categories	Lifetime Risk			Risk to age 50			Ref Fig.
			RF	PRS	RF and PRS	RF	PRS	RF and PRS	
Untested	NA	population	99.9	100.0	99.4	100.0	100.0	100.0	2(a)(c) and s2(a)(c)
		moderate	0.1	0.0	0.6	0.0	0.0	0.0	
		high	0.0	0.0	0.0	0.0	0.0	0.0	
	M 50	population	59.1	61.8	61.7	100.0	100.0	100.0	2(b)(d) and s2(b)(d)
		moderate	40.3	38.2	37.1	0.0	0.0	0.0	
		high	0.6	0.0	1.2	0.0	0.0	0.0	
BRCA1	NA	population	0.0	0.0	0.0	0.0	0.0	0.3	3(a) and s3(a)
		moderate	0.0	0.0	0.0	1.5	0.1	4.1	
		high	100.0	100.0	100.0	98.5	99.9	95.6	
	M 50	population	0.0	0.0	0.0	0.0	0.0	0.0	3(b) and s3(b)
		moderate	0.0	0.0	0.0	0.0	0.0	0.2	
		high	100.0	100.0	100.0	100.0	100.0	99.8	
BRCA2	NA	population	1.6	0.1	4.7	100.0	100.0	99.9	3(c) and s3(c)
		moderate	42.8	37.7	42.3	0.0	0.0	0.1	
		high	55.6	62.2	52.9	0.0	0.0	0.0	
	M 50	population	0.0	0.0	0.1	95.9	99.9	95.3	3(d) and s3(d)
		moderate	1.9	0.1	4.7	4.0	0.1	4.5	
		high	98.1	99.9	95.3	0.0	0.0	0.2	

PV	Family	Risk	Lifetime Risk			Risk to age 50			Ref Fig.
RAD51D	NA	population	0.6	0.0	2.1	100.0	100.0	99.8	3(e) and s3(e)
		moderate	24.5	18.3	30.8	0.0	0.0	0.2	
		high	75.0	81.7	67.1	0.0	0.0	0.0	
	M 50	population	0.0	0.0	0.0	93.5	99.6	92.6	3(f) and s3(f)
		moderate	0.6	0.0	2.0	6.3	0.4	7.0	
		high	99.4	100.0	98.0	0.2	0.0	0.4	
RAD51C	NA	population	1.8	0.1	4.9	99.8	100.0	99.3	3(g) and s3(g)
		moderate	44.4	39.3	43.1	0.2	0.0	0.7	
		high	53.7	60.6	52.0	0.0	0.0	0.0	
	M 50	population	0.0	0.0	0.1	85.8	96.2	85.0	3(h) and s3(h)
		moderate	2.0	0.1	5.0	13.6	3.8	13.7	
		high	98.0	99.9	95.0	0.6	0.0	1.3	
BRIP1	NA	population	43.6	34.0	45.7	100.0	100.0	99.7	3(i) and s3(i)
		moderate	52.3	65.0	47.4	0.0	0.0	0.3	
		high	4.1	1.0	6.9	0.0	0.0	0.0	
	M 50	population	2.1	0.1	4.8	93.0	99.3	91.4	3(j) and s3(j)
		moderate	43.5	38.9	44.5	6.8	0.7	8.1	
		high	54.5	61.1	50.7	0.3	0.0	0.5	

Figure 3 shows the predicted lifetime EOC risk for carriers of PVs in *BRCA1, BRCA2, RAD51D*, *RAD51C* and *BRIP1* based on RFs and PRS for two FH scenarios. Taking a *RAD51D* PV carrier, for example, based on PV testing and FH alone, the predicted risks are 13% when FH is unknown and 23% when having a mother diagnosed with EOC at age 50. When RFs and the PRS are considered jointly, risks vary from 4% for those at the 1st percentile to 28% for the 99th with unknown FH and from 9% to 43% with an affected mother. Table 1 shows the proportion of females with PVs falling into different risk categories. Based on the combined distribution, 33% of *RAD51D* PV carriers in the population are expected to have a lifetime EOC risk of less than 10%. Similarly, the distributions of risk for *BRIP*1 PV carriers are shown in Figure 3(i) and 3(j) and in Table 1. Based on the combined RFs and PRS distributions, 46% of *BRIP1* PV carriers in the population are expected to have lifetime risks of less than 5%; 47% to have risks between 5% and 10%, and 7% to have risks of 10% or greater. A *BRIP1* PV

carrier with an affected mother, on the basis of FH alone, has a lifetime risk of 11%. However, when the RFs and PRS are considered, 50% of those would be reclassified as having lifetime risks of less than 10%.

Figures s4 and s5 show the probability trees describing the reclassification of females as more information (RFs, PRS, and testing for PVs in the MGs) is added to the model for a female with unknown FH and a female with a mother diagnosed at age 50 respectively based on the predicted lifetime risks. Figures s4(a) and s5(a) show the reclassification resulting from adding RFs, MG and PRS sequentially, while Figures s4(b) and s5(b) assume the order RFs, PRS and then MG. Assuming the three risk categories for lifetime risks are <5%, ≥5% but <10% and ≥10%, there is significant reclassification as more information is added.

Model validation

After censoring, 1961 participants with 374 incident cases and 1587 controls met the 5-year risk prediction eligibility criteria. Table s5 summarises their characteristics at baseline.

The model considering FH, the 15-variant PRS and a subset of the RFs (but not including testing for PVs in the MGs) demonstrated good calibration in both absolute and relative predicted risk (Figure 4). Over the five-year period, the model predicted 391 EOCs, close to the 374 observed (E/O=1.05, 95%CI=0.94-1.16). The model was well calibrated across the quintiles of predicted risk (HL p=0.08), although there was a suggestion for an underprediction of risk in the lowest quintile (absolute risk E/O=0.66, 95%CI=0.52-0.91; RR E/O=0.63, 95%CI:0.42-0.95). The AUC for assessing discrimination of these model components was 0.61 (95%CI:0.58-0.64).

When looking at individual factors, FH predicted the widest 5-year risk variability (sd=0.0013; range 0.04% to 4.0%), followed by RFs (sd=0.0010; range:0.02%-0.7%) and PRS (sd=0.0009; range:0.05%-1.0%, Figure s6). As expected, their sequential inclusion increased the variability (sd=0.0018; Figure s6).

DISCUSSION

The EOC risk prediction model presented here combines the effects of FH, the explicit effects of rare moderate- to high-risk PVs in five established EOC susceptibility genes, a 36-variant PRS and other clinical and epidemiological factors (Table 1). The model provides a consistent approach for estimating EOC risk on the basis of all known factors and allows for prevention approaches to be targeted at those at highest risk.

The results demonstrate that in the general population (unknown FH), the existing PRS and RF alone identify 0.6% of females who have a lifetime risk of >5% (Table 2). On the other hand, for females with FH, 37.1% of females would have a predicted risk between 5% and 10% and 1.2% would have an EOC risk of ≥10% (Table 2). The results show that the RFs provide a somewhat greater level of risk stratification than the 36-variant PRS. However, discrimination is greater when both are considered jointly. These results were in line with the observed risk distributions in the validation dataset, but direct comparisons were not possible due to the different variants included in the PRSs and limited RFs in the validation study. The results also show that significant levels of risk re-categorisation can occur for carriers of PVs in moderate or high-risk susceptibility genes.

The comprehensive risk model is based on a synthetic approach previously used for BC[28] and makes several assumptions. In particular, we assumed that the risks associated with known RFs and the PRS combine multiplicatively. We have not assessed this assumption in

the present study, however, published studies found no evidence of deviations from the multiplicative model for the combined effect of the RFs and the PRS,[26] suggesting that this assumption is reasonable. The model assumes that the RFs are also independent of the residual polygenic component that captures the effect of FH. However, for the RFs included, we used estimates from published studies that have adjusted for the other known EOC RFs. The observation that the model was calibrated on the RR scale in the UKCTOCS validation study also suggests that these assumptions are broadly valid.

Similarly, the model assumes that the relative effect-sizes of RFs and the PRS are similar in females carrying PVs in *BRCA1, BRCA2, RAD51C, RAD51D* and *BRIP1* and those without PVs in these genes. Evidence from studies of *BRCA1* and *BRCA2* PV carriers suggests that this assumption is plausible: PRSs for EOC have been shown to be associated with similar RRs in the general population and in *BRCA1* and *BRCA2* PV carriers.[32, 36, 37] The current evidence also suggests that known RFs have similar effect sizes in *BRCA1* and *BRCA2* PV carriers as in non-carriers.[38, 39] No studies have so far assessed the joint effects of *RAD51C, RAD51D* and *BRIP1* PVs with the PRS, but the observation that FH modifies EOC risk for *RAD51C/D* PV carriers[27] suggests that similar arguments are likely to apply. Large prospective studies are required to address these questions in more detail. We were not able to validate these assumptions explicitly in UKCTOCS because gene-panel testing data were not available.

Other RFs for EOC that have been reported in the literature include breast feeding[40] and age at menarche and menopause.[23] However, the evidence for these RFs is still limited. Our model is flexible enough to allow for additional RFs to be incorporated in the future.

We validated the five-year predicted risks on the basis of FH, RFs and PRS available in an independent dataset from a prospective trial.[2] A key strength was that EOC was a primary outcome in UKCTOCS. All cases were reviewed and confirmed by an independent outcome review committee.[2] The results indicated that absolute and RRs were well-calibrated overall and in the top quintiles of predicted risk. However, there was some underprediction of EOC in the bottom quintile. This could be due to differences in the RF distributions in those who volunteer to participate in research (self-selected more healthy individuals[41]) compared to the general population or due to random variations in the effects of the RFs in UKCTOCS compared to other studies. Alternatively, the multiplicative assumption may break down in the lowest risk category. Further large prospective cohorts will be required to determine whether the underprediction in the lowest risk category reflects a systematic miscalibration of the model or is due to chance. Although the AUC based on model components in this validation study was modest, it is not surprising given that only a subset of the model predictors were used, and UKCTOCS recruited primarily low-risk women. Inclusion of the optimal PRS,[32] all RFs and information on PVs in the five genes that account for a large fraction of the EOC FRR are expected to lead to an increase in AUC.

The current validation study has some limitations. The underlying model accounts for FH information on both affected and unaffected family members, but the UKCTOCS recruitment questionnaire did not include information on unaffected family members. Family sizes and ages for unobserved family members were imputed using demographic data. In addition, since data on whether the affected family members were from the paternal or maternal side, we assumed all the affected family members were from the same (maternal) side. These may result in inaccuracies in risk predictions. A further limitation is that UKCTOCS was undertaken to assess screening of low-risk women and therefore is not

necessarily representative of a true population cohort, as females with a FH of two or more relatives with EOC or who were known carriers of *BRCA1/2* PVs were not eligible to participate in the randomised controlled trial. Data were not available on the rare moderate and high-risk PVs, and we were only able to assess a PRS with 15-variants, rather than the more informative 36-variant PRS. Therefore, it has not been possible to validate the full model presented here. Future analyses in other cohorts will be required to further validate the full model.

In summary, we have presented a methodological framework for a comprehensive EOC risk prediction model that considers the currently known genetic and epidemiologic RFs and explicit FH. The model allows users to obtain consistent, individualised EOC risks. It can also be used to identify target populations for studies to assess novel prevention strategies (such as salpingectomy) or early detection approaches by identifying those at higher risk of developing the disease for enrolment into such studies. Future independent studies should aim to validate the full model, including the full PRS and rare PVs in diverse settings. The model is available via the CanRisk Tool (www.canrisk.org), a user-friendly web tool that allows users to obtain future risks of developing EOC.

Acknowledgements

This work has been supported by grants from Cancer Research UK (C12292/A20861).

The analysis is part of PROMISE, which was funded through Cancer Research UK PRC Programme Grant A12677 and by The Eve Appeal. University College London investigators received support from the National Institute for Health Research University College London Hospitals Biomedical Research Centre and from MRC core funding (MR_UU_12023). UKCTOCS was core funded by the Medical Research Council (G9901012 and G0801228), Cancer Research UK (C1479/A2884), and the Department of Health with additional support from the Eve Appeal.

FMW was supported by a National Institute for Health Research (NIHR) Clinician Scientist award (NIHR-CS-012-03).

MT was funded by the European Union Seventh Framework Program (2007–2013)/European Research Council (310018).

The authors are particularly grateful to those throughout the UK who are participating in the trial and to the centre leads and the entire medical, nursing and administrative staff who work on the UKCTOCS.

Ethics Declaration

Ethics approval UKCTOCS was approved by the UK North West Multicentre Research Ethics Committees (North West MREC 00/8/34) on 21 June 2000 with site-specific approval from the local regional ethics committees and the Caldicott guardians (data controllers) of the primary care trusts. The SNP protocol was approved by NRES Committee North West - Liverpool Central (14/NW1026) in June 2014.

Conflicts of Interest

The authors Douglas F. Easton, Antonis C. Antoniou, Alex P. Cunningham, Andrew Lee and Tim Carver are listed as creators of the BOADICEA model, which has been licensed to Cambridge Enterprise for commercialisation.

Usha Menon has shares in Abcodia awarded to her by UCL.

Data availability

The model is freely available at www.canrisk.org. For access to UKCTOCS data set, which is subject to GDPR rules, please contact the UKCTOCS Biobank coordinator (s.apostolidou@ucl.ac.uk). The data access process is outlined here http://uklwc.mrcctu.ucl.ac.uk/access-process/

Author Contributions

Conceptualization: ACA, PPDP, DFE

Methodology: AL, XY, ACA, DFE, PPDP, JT

Software: AL, XY, APC, TC, SA

Investigation: AL, XY, ACA, PPDP, DFE, SAG, SJR, FG, RM, FWM, MT

Data curation: AGM, AR, GL

Formal analysis: AL, XY, JT, NM

Funding acquisition: ACA, IJ, UM

Resources: PPDP, AGM, AR, JT, SAG, SJR, IJ, UM

Supervision: ACA

Visualisation: AL, XY

Writing – original draft: AL, XY, ACA

Writing – review & editing: all authors

REFERENCES

1. Cancer Research UK. Ovarian Cancer Statistics [Ovarian cancer statistics for the UK]. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/ovarian-cancer.

2. Jacobs IJ, Menon U, Ryan A, Gentry-Maharaj A, Burnell M, Kalsi JK, Amso NN, Apostolidou S, Benjamin E, Cruickshank D, Crump DN, Davies SK, Dawnay A, Dobbs S, Fletcher G, Ford J, Godfrey K, Gunu R, Habib M, Hallett R, Herod J, Jenkins H, Karpinskyj C, Leeson S, Lewis SJ, Liston WR, Lopes A, Mould T, Murdoch J, Oram D, Rabideau DJ, Reynolds K, Scott I, Seif MW, Sharma A, Singh N, Taylor J, Warburton F, Widschwendter M, Williamson K, Woolas R, Fallowfield L, McGuire AJ, Campbell S, Parmar M, Skates SJ. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet*. 2016;387(10022):945-56.

3. Rosenthal AN, Fraser LSM, Philpott S, Manchanda R, Burnell M, Badman P, Hadwin R, Rizzuto I, Benjamin E, Singh N, Evans DG, Eccles DM, Ryan A, Liston R, Dawnay A, Ford J, Gunu R, Mackay J, Skates SJ, Menon U, Jacobs IJ, United Kingdom Familial Ovarian Cancer Screening Study c. Evidence of Stage Shift in Women Diagnosed With Ovarian Cancer During Phase II of the United Kingdom Familial Ovarian Cancer Screening Study. *J Clin Oncol*. 2017;35(13):1411-20.

4. Ludwig KK, Neuner J, Butler A, Geurts JL, Kong AL. Risk reduction and survival benefit of prophylactic surgery in BRCA mutation carriers, a systematic review. *Am J Surg*. 2016;212(4):660-9.

5. Manchanda R, Legood R, Antoniou AC, Gordeev VS, Menon U. Specifying the ovarian cancer risk threshold of 'premenopausal risk-reducing salpingo-oophorectomy' for ovarian cancer prevention: a cost-effectiveness analysis. *J Med Genet*. 2016;53(9):591-9.

6. Manchanda R, Legood R, Pearce L, Menon U. Defining the risk threshold for risk reducing salpingo-oophorectomy for ovarian cancer prevention in low risk postmenopausal women. *Gynecol Oncol*. 2015;139(3):487-94.

7. Parker WH, Feskanich D, Broder MS, Chang E, Shoupe D, Farquhar CM, Berek JS, Manson JE. Long-term mortality associated with oophorectomy compared with ovarian conservation in the nurses' health study. *Obstet Gynecol*. 2013;121(4):709-16.

8. Rocca WA, Gazzuola-Rocca L, Smith CY, Grossardt BR, Faubion SS, Shuster LT, Kirkland JL, Stewart EA, Miller VM. Accelerated Accumulation of Multimorbidity After Bilateral Oophorectomy: A Population-Based Cohort Study. *Mayo Clin Proc*. 2016;91(11):1577-89.

9. Parker WH, Jacoby V, Shoupe D, Rocca W. Effect of bilateral oophorectomy on women's long-term health. *Womens Health (Lond)*. 2009;5(5):565-76.

10. Gaba F, Manchanda R. Systematic review of acceptability, cardiovascular, neurological, bone health and HRT outcomes following risk reducing surgery in BRCA carriers. *Best Pract Res Clin Obstet Gynaecol*. 2020;65:46-65.

11. Jervis S, Song H, Lee A, Dicks E, Harrington P, Baynes C, Manchanda R, Easton DF, Jacobs I, Pharoah PP, Antoniou AC. A risk prediction algorithm for ovarian cancer incorporating BRCA1, BRCA2, common alleles and other familial effects. *J Med Genet*. 2015;52(7):465-75.

12. Loveday C, Turnbull C, Ramsay E, Hughes D, Ruark E, Frankum JR, Bowden G, Kalmyrzaev B, Warren-Perry M, Snape K, Adlard JW, Barwell J, Berg J, Brady AF, Brewer C, Brice G, Chapman C, Cook J, Davidson R, Donaldson A, Douglas F, Greenhalgh L, Henderson A, Izatt L, Kumar A, Lalloo F, Miedzybrodzka Z, Morrison PJ, Paterson J, Porteous M, Rogers MT, Shanley S, Walker L, Breast Cancer Susceptibility C, Eccles D, Evans DG, Renwick A, Seal

S, Lord CJ, Ashworth A, Reis-Filho JS, Antoniou AC, Rahman N. Germline mutations in RAD51D confer susceptibility to ovarian cancer. *Nat Genet*. 2011;43(9):879-82.

13. Loveday C, Turnbull C, Ruark E, Xicola RM, Ramsay E, Hughes D, Warren-Perry M, Snape K, Breast Cancer Susceptibility C, Eccles D, Evans DG, Gore M, Renwick A, Seal S, Antoniou AC, Rahman N. Germline RAD51C mutations confer susceptibility to ovarian cancer. *Nat Genet*. 2012;44(5):475-6; author reply 6.

14. Ramus SJ, Song H, Dicks E, Tyrer JP, Rosenthal AN, Intermaggio MP, Fraser L, Gentry-Maharaj A, Hayward J, Philpott S, Anderson C, Edlund CK, Conti D, Harrington P, Barrowdale D, Bowtell DD, Alsop K, Mitchell G, Group AS, Cicek MS, Cunningham JM, Fridley BL, Alsop J, Jimenez-Linan M, Poblete S, Lele S, Sucheston-Campbell L, Moysich KB, Sieh W, McGuire V, Lester J, Bogdanova N, Durst M, Hillemanns P, Ovarian Cancer Association C, Odunsi K, Whittemore AS, Karlan BY, Dork T, Goode EL, Menon U, Jacobs IJ, Antoniou AC, Pharoah PD, Gayther SA. Germline Mutations in the BRIP1, BARD1, PALB2, and NBN Genes in Women With Ovarian Cancer. *J Natl Cancer Inst*. 2015;107(11).

15. Pal T, Akbari MR, Sun P, Lee JH, Fulp J, Thompson Z, Coppola D, Nicosia S, Sellers TA, McLaughlin J, Risch HA, Rosen B, Shaw P, Schildkraut J, Narod SA. Frequency of mutations in mismatch repair genes in a population-based study of women with ovarian cancer. *Br J Cancer*. 2012;107(10):1783-90.

16. Yang X, Leslie G, Doroszuk A, Schneider S, Allen J, Decker B, Dunning AM, Redman J, Scarth J, Plaskocinska I, Luccarini C, Shah M, Pooley K, Dorling L, Lee A, Adank MA, Adlard J, Aittomaki K, Andrulis IL, Ang P, Barwell J, Bernstein JL, Bobolis K, Borg A, Blomqvist C, Claes KBM, Concannon P, Cuggia A, Culver JO, Damiola F, de Pauw A, Diez O, Dolinsky JS, Domchek SM, Engel C, Evans DG, Fostira F, Garber J, Golmard L, Goode EL, Gruber SB, Hahnen E, Hake C, Heikkinen T, Hurley JE, Janavicius R, Kleibl Z, Kleiblova P, Konstantopoulou I, Kvist A,

Laduca H, Lee ASG, Lesueur F, Maher ER, Mannermaa A, Manoukian S, McFarland R, McKinnon W, Meindl A, Metcalfe K, Mohd Taib NA, Moilanen J, Nathanson KL, Neuhausen S, Ng PS, Nguyen-Dumont T, Nielsen SM, Obermair F, Offit K, Olopade OI, Ottini L, Penkert J, Pylkas K, Radice P, Ramus SJ, Rudaitis V, Side L, Silva-Smith R, Silvestri V, Skytte AB, Slavin T, Soukupova J, Tondini C, Trainer AH, Unzeitig G, Usha L, van Overeem Hansen T, Whitworth J, Wood M, Yip CH, Yoon SY, Yussuf A, Zogopoulos G, Goldgar D, Hopper JL, Chenevix-Trench G, Pharoah P, George SHL, Balmana J, Houdayer C, James P, El-Haffaf Z, Ehrencrona H, Janatova M, Peterlongo P, Nevanlinna H, Schmutzler R, Teo SH, Robson M, Pal T, Couch F, Weitzel JN, Elliott A, Southey M, Winqvist R, Easton DF, Foulkes WD, Antoniou AC, Tischkowitz M. Cancer Risks Associated With Germline PALB2 Pathogenic Variants: An International Study of 524 Families. *J Clin Oncol*. 2020;38(7):674-85.

17. Kuchenbaecker KB, Ramus SJ, Tyrer J, Lee A, Shen HC, Beesley J, Lawrenson K, McGuffog L, Healey S, Lee JM, Spindler TJ, Lin YG, Pejovic T, Bean Y, Li Q, Coetzee S, Hazelett D, Miron A, Southey M, Terry MB, Goldgar DE, Buys SS, Janavicius R, Dorfling CM, van Rensburg EJ, Neuhausen SL, Ding YC, Hansen TVO, Jønson L, Gerdes A-M, Ejlertsen B, Barrowdale D, Dennis J, Benitez J, Osorio A, Garcia MJ, Komenaka I, Weitzel JN, Ganschow P, Peterlongo P, Bernard L, Viel A, Bonanni B, Peissel B, Manoukian S, Radice P, Papi L, Ottini L, Fostira F, Konstantopoulou I, Garber J, Frost D, Perkins J, Platte R, Ellis S, Godwin AK, Schmutzler RK, Meindl A, Engel C, Sutter C, Sinilnikova OM, Damiola F, Mazoyer S, Stoppa-Lyonnet D, Claes K, De Leeneer K, Kirk J, Rodriguez GC, Piedmonte M, O'Malley DM, de la Hoya M, Caldes T, Aittomäki K, Nevanlinna H, Collée JM, Rookus MA, Oosterwijk JC, Tihomirova L, Tung N, Hamann U, Isaccs C, Tischkowitz M, Imyanitov EN, Caligo MA, Campbell IG, Hogervorst FBL, Olah E, Diez O, Blanco I, Brunet J, Lazaro C, Pujana MA, Jakubowska A, Gronwald J, Lubinski J, Sukiennicki G, Barkardottir RB, Plante M, Simard J,

Soucy P, Montagna M, Tognazzo S, Teixeira MR, Pankratz VS, Wang X, Lindor N, Szabo CI, Kauff N, Vijai J, Aghajanian CA, Pfeiler G, Berger A, Singer CF, Tea M-K, Phelan CM, Greene MH, Mai PL, Rennert G, Mulligan AM, Tchatchou S, Andrulis IL, Glendon G, Toland AE, Jensen UB, Kruse TA, Thomassen M, Bojesen A, Zidan J, Friedman E, Laitman Y, Soller M, Liljegren A, Arver B, Einbeigi Z, Stenmark-Askmalm M, Olopade OI, Nussbaum RL, Rebbeck TR, Nathanson KL, Domchek SM, Lu KH, Karlan BY, Walsh C, Lester J, Hein A, Ekici AB, Beckmann MW, Fasching PA, Lambrechts D, Van Nieuwenhuysen E, Vergote I, Lambrechts S, Dicks E, Doherty JA, Wicklund KG, Rossing MA, Rudolph A, Chang-Claude J, Wang-Gohrke S, Eilber U, Moysich KB, Odunsi K, Sucheston L, Lele S, Wilkens LR, Goodman MT, Thompson PJ, Shvetsov YB, Runnebaum IB, Dürst M, Hillemanns P, Dörk T, Antonenkova N, Bogdanova N, Leminen A, Pelttari LM, Butzow R, Modugno F, Kelley JL, Edwards RP, Ness RB, du Bois A, Heitz F, Schwaab I, Harter P, Matsuo K, Hosono S, Orsulic S, Jensen A, Kjaer SK, Hogdall E, Hasmad HN, Azmi MAN, Teo S-H, Woo Y-L, Fridley BL, Goode EL, Cunningham JM, Vierkant RA, Bruinsma F, Giles GG, Liang D, Hildebrandt MAT, Wu X, Levine DA, Bisogna M, Berchuck A, Iversen ES, Schildkraut JM, Concannon P, Weber RP, Cramer DW, Terry KL, Poole EM, Tworoger SS, Bandera EV, Orlow I, Olson SH, Krakstad C, Salvesen HB, Tangen IL, Bjorge L, van Altena AM, Aben KKH, Kiemeney LA, Massuger LFAG, Kellar M, Brooks-Wilson A, Kelemen LE, Cook LS, Le ND, Cybulski C, Yang H, Lissowska J, Brinton LA, Wentzensen N, Hogdall C, Lundvall L, Nedergaard L, Baker H, Song H, Eccles D, McNeish I, Paul J, Carty K, Siddiqui N, Glasspool R, Whittemore AS, Rothstein JH, McGuire V, Sieh W, Ji B-T, Zheng W, Shu X-O, Gao Y-T, Rosen B, Risch HA, McLaughlin JR, Narod SA, Monteiro AN, Chen A, Lin H-Y, Permuth-Wey J, Sellers TA, Tsai Y-Y, Chen Z, Ziogas A, Anton-Culver H, Gentry-Maharaj A, Menon U, Harrington P, Lee AW, Wu AH, Pearce CL, Coetzee G, Pike MC, Dansonka-Mieszkowska A, Timorek A, Rzepecka IK, Kupryjanczyk J, Freedman M, Noushmehr H, Easton

DF, Offit K, Couch FJ, Gayther S, Pharoah PP, Antoniou AC, Chenevix-Trench G, Easton DF, Offit K, Couch FJ, Gayther S, Pharoah PP, Antoniou AC, Chenevix-Trench G, Brca tColoMoB. Identification of six new susceptibility loci for invasive epithelial ovarian cancer. *Nature Genetics*. 2015;47(2):164-71.

18. Phelan CM, Kuchenbaecker KB, Tyrer JP, Kar SP, Lawrenson K, Winham SJ, Dennis J, Pirie A, Riggan MJ, Chornokur G, Earp MA, Lyra PC, Jr., Lee JM, Coetzee S, Beesley J, McGuffog L, Soucy P, Dicks E, Lee A, Barrowdale D, Lecarpentier J, Leslie G, Aalfs CM, Aben KK, Adams M, Adlard J, Andrulis IL, Anton-Culver H, Antonenkova N, group As, Aravantinos G, Arnold N, Arun BK, Arver B, Azzollini J, Balmana J, Banerjee SN, Barjhoux L, Barkardottir RB, Bean Y, Beckmann MW, Beeghly-Fadiel A, Benitez J, Bermisheva M, Bernardini MQ, Birrer MJ, Bjorge L, Black A, Blankstein K, Blok MJ, Bodelon C, Bogdanova N, Bojesen A, Bonanni B, Borg A, Bradbury AR, Brenton JD, Brewer C, Brinton L, Broberg P, Brooks-Wilson A, Bruinsma F, Brunet J, Buecher B, Butzow R, Buys SS, Caldes T, Caligo MA, Campbell I, Cannioto R, Carney ME, Cescon T, Chan SB, Chang-Claude J, Chanock S, Chen XQ, Chiew YE, Chiquette J, Chung WK, Claes KB, Conner T, Cook LS, Cook J, Cramer DW, Cunningham JM, D'Aloisio AA, Daly MB, Damiola F, Damirovna SD, Dansonka-Mieszkowska A, Dao F, Davidson R, DeFazio A, Delnatte C, Doheny KF, Diez O, Ding YC, Doherty JA, Domchek SM, Dorfling CM, Dork T, Dossus L, Duran M, Durst M, Dworniczak B, Eccles D, Edwards T, Eeles R, Eilber U, Ejlertsen B, Ekici AB, Ellis S, Elvira M, Study E, Eng KH, Engel C, Evans DG, Fasching PA, Ferguson S, Ferrer SF, Flanagan JM, Fogarty ZC, Fortner RT, Fostira F, Foulkes WD, Fountzilas G, Fridley BL, Friebel TM, Friedman E, Frost D, Ganz PA, Garber J, Garcia MJ, Garcia-Barberan V, Gehrig A, Collaborators GS, Gentry-Maharaj A, Gerdes AM, Giles GG, Glasspool R, Glendon G, Godwin AK, Goldgar DE, Goranova T, Gore M, Greene MH, Gronwald J, Gruber S, Hahnen E, Haiman CA, Hakansson N, Hamann U, Hansen TV, Harrington PA, Harris HR,

Hauke J, Study H, Hein A, Henderson A, Hildebrandt MA, Hillemanns P, Hodgson S, Hogdall CK, Hogdall E, Hogervorst FB, Holland H, Hooning MJ, Hosking K, Huang RY, Hulick PJ, Hung J, Hunter DJ, Huntsman DG, Huzarski T, Imyanitov EN, Isaacs C, Iversen ES, Izatt L, Izquierdo A, Jakubowska A, James P, Janavicius R, Jernetz M, Jensen A, Jensen UB, John EM, Johnatty S, Jones ME, Kannisto P, Karlan BY, Karnezis A, Kast K, Investigators KC, Kennedy CJ, Khusnutdinova E, Kiemeney LA, Kiiski JI, Kim SW, Kjaer SK, Kobel M, Kopperud RK, Kruse TA, Kupryjanczyk J, Kwong A, Laitman Y, Lambrechts D, Larranaga N, Larson MC, Lazaro C, Le ND, Le Marchand L, Lee JW, Lele SB, Leminen A, Leroux D, Lester J, Lesueur F, Levine DA, Liang D, Liebrich C, Lilyquist J, Lipworth L, Lissowska J, Lu KH, Lubinnski J, Luccarini C, Lundvall L, Mai PL, Mendoza-Fandino G, Manoukian S, Massuger LF, May T, Mazoyer S, McAlpine JN, McGuire V, McLaughlin JR, McNeish I, Meijers-Heijboer H, Meindl A, Menon U, Mensenkamp AR, Merritt MA, Milne RL, Mitchell G, Modugno F, Moes-Sosnowska J, Moffitt M, Montagna M, Moysich KB, Mulligan AM, Musinsky J, Nathanson KL, Nedergaard L, Ness RB, Neuhausen SL, Nevanlinna H, Niederacher D, Nussbaum RL, Odunsi K, Olah E, Olopade OI, Olsson H, Olswold C, O'Malley DM, Ong KR, Onland-Moret NC, group Os, Orr N, Orsulic S, Osorio A, Palli D, Papi L, Park-Simon TW, Paul J, Pearce CL, Pedersen IS, Peeters PH, Peissel B, Peixoto A, Pejovic T, Pelttari LM, Permuth JB, Peterlongo P, Pezzani L, Pfeiler G, Phillips KA, Piedmonte M, Pike MC, Piskorz AM, Poblete SR, Pocza T, Poole EM, Poppe B, Porteous ME, Prieur F, Prokofyeva D, Pugh E, Pujana MA, Pujol P, Radice P, Rantala J, Rappaport-Fuerhauser C, Rennert G, Rhiem K, Rice P, Richardson A, Robson M, Rodriguez GC, Rodriguez-Antona C, Romm J, Rookus MA, Rossing MA, Rothstein JH, Rudolph A, Runnebaum IB, Salvesen HB, Sandler DP, Schoemaker MJ, Senter L, Setiawan VW, Severi G, Sharma P, Shelford T, Siddiqui N, Side LE, Sieh W, Singer CF, Sobol H, Song H, Southey MC, Spurdle AB, Stadler Z, Steinemann D, Stoppa-Lyonnet D, Sucheston-Campbell LE, Sukiennicki

G, Sutphen R, Sutter C, Swerdlow AJ, Szabo CI, Szafron L, Tan YY, Taylor JA, Tea MK, Teixeira MR, Teo SH, Terry KL, Thompson PJ, Thomsen LC, Thull DL, Tihomirova L, Tinker AV, Tischkowitz M, Tognazzo S, Toland AE, Tone A, Trabert B, Travis RC, Trichopoulou A, Tung N, Tworoger SS, van Altena AM, Van Den Berg D, van der Hout AH, van der Luijt RB, Van Heetvelde M, Van Nieuwenhuysen E, van Rensburg EJ, Vanderstichele A, Varon-Mateeva R, Vega A, Edwards DV, Vergote I, Vierkant RA, Vijai J, Vratimos A, Walker L, Walsh C, Wand D, Wang-Gohrke S, Wappenschmidt B, Webb PM, Weinberg CR, Weitzel JN, Wentzensen N, Whittemore AS, Wijnen JT, Wilkens LR, Wolk A, Woo M, Wu X, Wu AH, Yang H, Yannoukakos D, Ziogas A, Zorn KK, Narod SA, Easton DF, Amos CI, Schildkraut JM, Ramus SJ, Ottini L, Goodman MT, Park SK, Kelemen LE, Risch HA, Thomassen M, Offit K, Simard J, Schmutzler RK, Hazelett D, Monteiro AN, Couch FJ, Berchuck A, Chenevix-Trench G, Goode EL, Sellers TA, Gayther SA, Antoniou AC, Pharoah PD. Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer. *Nat Genet*. 2017.

19. Collaborative Group on Epidemiological Studies of Ovarian Cancer. Ovarian cancer and body size: individual participant meta-analysis including 25,157 women with ovarian cancer from 47 epidemiological studies. *PLoS Med*. 2012;9(4):e1001200.

20. Cancer CGoESoO. Menopausal hormone use and ovarian cancer risk: individual participant meta-analysis of 52 epidemiological studies. *The Lancet*. 2015;385(9980):1835-42.

21. Li K, Husing A, Fortner RT, Tjonneland A, Hansen L, Dossus L, Chang-Claude J, Bergmann M, Steffen A, Bamia C, Trichopoulos D, Trichopoulou A, Palli D, Mattiello A, Agnoli C, Tumino R, Onland-Moret NC, Peeters PH, Bueno-de-Mesquita HB, Gram IT, Weiderpass E, Sanchez-Cantalejo E, Chirlaque MD, Duell EJ, Ardanaz E, Idahl A, Lundin E, Khaw KT, Travis RC, Merritt MA, Gunter MJ, Riboli E, Ferrari P, Terry K, Cramer D, Kaaks R. An epidemiologic

risk prediction model for ovarian cancer in Europe: the EPIC study. *Br J Cancer*. 2015;112(7):1257-65.

22. Pearce CL, Stram DADO, Ness RB, Stram DADO, Roman LD, Templeman C, Lee AW, Menon U, Fasching PA, McAlpine JN, Doherty JA, Huntsman DG, Wu AH, Berchuck A, Pike MC, Pharoah PDP, Modugno F, Schildkraut JM, Rossing MA, Huntsman DG, Wu AH, Berchuck A, Pike MC, Pharoah PDP. Population distribution of lifetime risk of ovarian cancer in the United States. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2015;24(4):671-6.

23. Rosner BA, Colditz GA, Webb PM, Hankinson SE. Mathematical models of ovarian cancer incidence. *Epidemiology (Cambridge, Mass)*. 2005;16(4):508-15.

24. Gaitskell K, Coffey K, Green J, Pirie K, Reeves GK, Ahmed AA, Barnes I, Beral V, Abbiss H, Abbott S, Alison R, Allen N, Armstrong M, Baker K, Balkwill A, Banks E, Barnes I, Beral V, Black J, Blanks R, Bradbury K, Brown A, Cairns B, Canfell K, Canoy D, Chadwick A, Crossley B, Crowe F, Ewart D, Ewart S, Fletcher L, Floud S, Gathani T, Gerrard L, Goodill A, Green J, Guiver L, Hozak M, Lingard I, Kan SW, Kirichek O, Langston N, Liu B, Moser K, Pirie K, Reeves G, Shaw K, Sherman E, Strange H, Sweetland S, Tipper S, Travis R, Trickett L, Wright L, Yang O, Young H, Banks E, Beral V, Carpenter L, Dezateux C, Green J, Patnick J, Peto R, Sudlow C. Tubal ligation and incidence of 26 site-specific cancers in the Million Women Study. *British journal of cancer*. 2016;114(9):1033-7.

25. Pfeiffer RM, Park Y, Kreimer AR, Lacey JV, Jr., Pee D, Greenlee RT, Buys SS, Hollenbeck A, Rosner B, Gail MH, Hartge P. Risk prediction for breast, endometrial, and ovarian cancer in white women aged 50 y or older: derivation and validation from population-based cohort studies. *PLoS Med*. 2013;10(7):e1001492.

26. Clyde MA, Weber RP, Iversen ES, Poole EM, Doherty JA, Goodman MT, Ness RB, Risch HA, Rossing MA, Terry KL, Wentzensen N, Whittemore AS, Anton-Culver H, Bandera EV, Berchuck A, Carney ME, Cramer DW, Cunningham JM, Cushing-Haugen KL, Edwards RP, Fridley BL, Goode EL, Lurie G, McGuire V, Modugno F, Moysich KB, Olson SH, Pearce CL, Pike MC, Rothstein JH, Sellers TA, Sieh W, Stram D, Thompson PJ, Vierkant RA, Wicklund KG, Wu AH, Ziogas A, Tworoger SS, Schildkraut JM. Original Contribution Risk Prediction for Epithelial Ovarian Cancer in 11 United States–Based Case-Control Studies: Incorporation of Epidemiologic Risk Factors and 17 Confirmed Genetic Loci. *American Journal of Epidemiology*. 2016.

27. Yang X, Song H, Leslie G, Engel C, Hahnen E, Auber B, Horvath J, Kast K, Niederacher D, Turnbull C, Houlston R, Hanson H, Loveday C, Dolinsky JS, LaDuca H, Ramus SJ, Menon U, Rosenthal AN, Jacobs I, Gayther SA, Dicks E, Nevanlinna H, Aittomaki K, Pelttari LM, Ehrencrona H, Borg A, Kvist A, Rivera B, Hansen TVO, Djursby M, Lee A, Dennis J, Bowtell DD, Traficante N, Diez O, Balmana J, Gruber SB, Chenevix-Trench G, kConFab I, Jensen A, Kjaer SK, Hogdall E, Castera L, Garber J, Janavicius R, Osorio A, Golmard L, Vega A, Couch FJ, Robson M, Gronwald J, Domchek SM, Culver JO, de la Hoya M, Easton DF, Foulkes WD, Tischkowitz M, Meindl A, Schmutzler RK, Pharoah PDP, Antoniou AC. Ovarian and breast cancer risks associated with pathogenic variants in RAD51C and RAD51D. *J Natl Cancer Inst*. 2020.

28. Lee A, Mavaddat N, Wilcox AN, Cunningham AP, Carver T, Hartley S, Babb de Villiers C, Izquierdo A, Simard J, Schmidt MK, Walter FM, Chatterjee N, Garcia-Closas M, Tischkowitz M, Pharoah P, Easton DF, Antoniou AC. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genet Med*. 2019;21(8):1708-18.

29. Pal Choudhury P, Maas P, Wilcox A, Wheeler W, Brook M, Check D, Garcia-Closas M, Chatterjee N. iCARE: An R package to build, validate and apply absolute risk models. *PLoS One*. 2020;15(2):e0228198.

30. Pal Choudhury P, Brook MN, Hurson AN, Lee A, Mulder CV, Coulson P, Schoemaker MJ, Jones ME, Swerdlow AJ, Chatterjee N, Antoniou AC, Garcia-Closas M. Comparative validation of the BOADICEA and Tyrer-Cuzick breast cancer risk models incorporating classical risk factors and polygenic risk in a population-based prospective cohort of women of European ancestry. *Breast Cancer Res*. 2021;23(1):22.

31. Lee A, Cunningham AP, Tischkowitz M, Simard J, Pharoah PD, Easton DF, Antoniou AC. Incorporating truncating variants in PALB2, CHEK2, and ATM into the BOADICEA breast cancer risk model. *Genet Med*. 2016;18(12):1190-8.

32. Dareng EO, Tyrer J, Barnes DR, Jones MR, Yang X, Aben KKH, Adank MA, Agata SA, Andrulis IL, Anton-Culver H, Antonenkova NN, Aravantinos G, Arun BK, Augustinsson A, Balmaña J, Bandera EV, Barkardottir RB, Barrowdale D, Beckmann MW, Beeghly-Fadiel A, Benitez J, Bermisheva M, Bernardini MQ, Bjorge L, Black A, Bogdanova NV, Bonanni B, Borg A, Brenton J, Budzilowska A, Butzow R, Buys SS, Cai H, Caligo MA, Campbell I, Cannioto R, Cassingham H, Chang-Claude J, Chanock SJ, Chen K, Chiew Y-E, Chung WK, Claes KBM, Colanna S, Cook LS, Couch FJ, Daly MB, Dao F, Davies E, de la Hoya M, de Putter R, DePersia A, Devilee P, Diez O, Ding YC, Doherty JA, Domchek SM, Dörk T, du Bois A, Dürst M, Eccles DM, Eliassen HA, Engel C, Evans DG, Fasching P, Flanagan JM, Foretova L, Fortner RT, Friedman E, Ganz PA, Garber J, Gensini F, Giles GG, Glendon G, Godwin AK, Goodman MT, Greene MH, Gronwald J, Hahnen E, Haiman CA, Håkansson N, Hamann U, Hansen TVO, Harris HR, Hartman M, Heitz F, Hildebrandt MAT, Høgdall E, Høgdall CK, Hopper JL, Huang R-Y, Huff C, Hulick PJ, Huntsman DG, Imyanitov EN, Isaacs C, Anna Jakubowska A, James P,

Janavicius R, Jensen A, Johannsson OT, John EM, Jones M, Kang D, Karlan BY, Karnezis A, Kelemen LE, Khusnutdinova E, Kiemeney LA, Kim B-G, Kjaer SK, Komenaka I, Kupryjanczyk J, Kurian AW, Kwong A, Lambrechts D, Larson MC, Lazaro C, Le ND, Leslie G, Lester J, Lesueur F, Levine D, Li L, Li J, Loud JT, Lu KH, Lubiński J, Machackova E, Mai PL, Manoukian S, Marks J, Kim Matsuno R, Matsuo K, May T, McGuffog L, McLaughlin JR, McNeish IA, Mebirouk N, Menon U, Miller A, Milne RL, Minlikeeva A, Modugno F, Montagna M, Moysich KB, Munro E, Nathanson KL, Neuhausen SL, Nevanlinna H, Ngeow Yuen Yie J, Roed Nielsen H, Nielsen FC, Nikitina-Zake L, Odunsi K, Offit K, Olah E, Olbrecht S, Olopade OI, Olson SH, Olsson H, Osorio A, Papi L, Park SK, Parsons MT, Pathak H, Pedersen IS, Peixoto A, Pejovic T, Perez-Segura P, Permuth JB, Peshkin B, Peterlongo P, Piskorz A, Prokofyeva D, Radice P, Rantala J, Riggan M, Risch HB, Rodriguez-Antona C, Ross EA, Rossing MA, Runnebaum I, Sandler DP, Santamariña M, Soucy P, Schmutzler RK, Setiawan VW, Shan K, Sieh W, Simard J, Singer CF, Sokolenko A, Song H, Southey M, Steed H, Stoppa-Lyonnet D, Sutphen R, Swerdlow AJ, Tan YY, Teixeira M, Teo SH, Terry KL, Terry MB, Thomassen M, Thompson PJ, Thomsen LC, Hull DL, Tischkowitz M, Titus L, Toland A, Torres D, Trabert B, Travis R, Tung N, Tworoger SS, Valen E, van Altena AM, van der Hout AH, Van Nieuwenhuysen E, van Rensburg EJ, Vega A, Velez Edwards D, Vierkant RA, Wang F, Wappenschmidt B, Webb PM, Weinberg CR, Weitzel JN, Wentzensen N, White E, Whittemore AS, Winham SJ, Wolk A, Woo Y-L, Wu AH, Yan L, Yannoukakos D, Zavaglia KM, Zheng W, Ziogas A, Zorn KK, Easton D, Lawrenson K, deFazio A, Sellers TA, Ramus SJ, Pearce C, Montiero AN, Cunningham JM, Goode EL, Schildkraut JM, Berchuck A, Chenevix-Trench G, Gayther SA, Antoniou AC, Pharoah PDP. Polygenic Risk Modelling for Prediction of Epithelial Ovarian Cancer Risk. *MedRxiv* 20219220 [Preprint]. December 2, 2020 [cited 2021 April 6] https://doi.org/10.1101/2020.11.30.20219220.

33. Yang X, Leslie G, Gentry-Maharaj A, Ryan A, Intermaggio M, Lee A, Kalsi JK, Tyrer J, Gaba F, Manchanda R, Pharoah PDP, Gayther SA, Ramus SJ, Jacobs I, Menon U, Antoniou AC. Evaluation of polygenic risk scores for ovarian cancer risk prediction in a prospective cohort study. *J Med Genet*. 2018;55(8):546-54.

34. Whittemore AS, Halpern J. Two-stage sampling designs for external validation of personal risk models. *Statistical Methods in Medical Research*. 2016;25(4):1313-29.

35. Terry MB, Liao Y, Whittemore AS, Leoce N, Buchsbaum R, Zeinomar N, Dite GS, Chung WK, Knight JA, Southey MC, Milne RL, Goldgar D, Giles GG, McLachlan SA, Friedlander ML, Weideman PC, Glendon G, Nesci S, Andrulis IL, John EM, Phillips KA, Daly MB, Buys SS, Hopper JL, MacInnis RJ. 10-year performance of four models of breast cancer risk: a validation study. *The Lancet Oncology*. 2019;20(4):504-17.

36. Barnes DR, Rookus MA, McGuffog L, Leslie G, Mooij TM, J. D, N. M, Adlard J, Ahmed M, Aittomaki K, Andrieu N, I. A, Arnold N, Investigators C, Schmutlzer R, Simard J, Easton DF, Chenevix-Trench G, Antoniou AC. Polygenic risk scores and breast and epithelial ovarian cancer risks for carriers of BRCA1 and BRCA2 pathogenic variants. *Genet Med* 2020;in Press.

37. Kuchenbaecker KB, McGuffog L, Barrowdale D, Lee A, Soucy P, Dennis J, Domchek SM, Robson M, Spurdle AB, Ramus SJ, Mavaddat N, Terry MB, Neuhausen SL, Schmutzler RK, Simard J, Pharoah PD, Offit K, Couch FJ, Chenevix-Trench G, Easton DF, Antoniou AC. Evaluation of Polygenic Risk Scores for Breast and Ovarian Cancer Risk Prediction in BRCA1 and BRCA2 Mutation Carriers. *J Natl Cancer Inst*. 2017;109(7).

38. Antoniou AC, Rookus M, Andrieu N, Brohet R, Chang-Claude J, Peock S, Cook M, Evans DG, Eeles R, Nogues C, Faivre L, Gesta P, van Leeuwen FE, Ausems MG, Osorio A, Caldes T, Simard J, Lubinski J, Gerdes AM, Olah E, Furhauser C, Olsson H, Arver B, Radice P, Easton DF, Goldgar DE. Reproductive and hormonal factors, and ovarian cancer risk for

BRCA1 and BRCA2 mutation carriers: results from the International BRCA1/2 Carrier Cohort Study. *Cancer Epidemiol Biomarkers Prev*. 2009;18(2):601-10.

39. Milne RL, Antoniou AC. Modifiers of breast and ovarian cancer risks for BRCA1 and BRCA2 mutation carriers. *Endocr Relat Cancer*. 2016;23(10):T69-84.

40. Jordan SJ, Cushing-Haugen KL, Wicklund KG, Doherty JA, Rossing MA. Breast-feeding and risk of epithelial ovarian cancer. *Cancer causes & control : CCC*. 2012;23(6):919-27.

41. Burnell M, Gentry-Maharaj A, Ryan A, Apostolidou S, Habib M, Kalsi J, Skates S, Parmar M, Seif MW, Amso NN, Godfrey K, Oram D, Herod J, Williamson K, Jenkins H, Mould T, Woolas R, Murdoch J, Dobbs S, Leeson S, Cruickshank D, Campbell S, Fallowfield L, Jacobs I, Menon U. Impact on mortality and cancer incidence rates of using random invitation from population registers for recruitment to trials. *Trials*. 2011;12:61.

Figure Legends

Figure 1. Predicted lifetime (age 20 to 80 years) EOC risk by PV and family history. Each Figure shows the risks assuming the female is untested, has no PVs, or carries a PV in *BRCA1*, *BRCA2*, *RAD51D*, *RAD51C* or *BRIP1*. Figure (a) assumes unknown family history. Figures (b)-(e) assume an increasing number of affected first-degree relatives, as indicated by the pedigree diagram inserts. Predictions are based on UK EOC population incidence.

Figure 2. Predicted lifetime (age 20 to 80 years) EOC risk for a female untested for PVs based on the different predictors of risk (risk factors (RFs) and PRS). Figures (a) and (c) show the risk for a female with an unknown family history (equivalent to the distribution of risk in the population), while figures (b) and (d) show the risk for a female with a mother affected at age 50. Figures (a) and (b) show the probability density function against absolute risk, while figures (c) and (d) show absolute risk against cumulative distribution. The vertical line in figure (a) and the horizontal line in Figure (c) (labelled "no RFs or PRS") is equivalent to the population risk of EOC. The "Population" risk is shown separately in figures (b) and (d). Predictions are based on UK EOC population incidences.

Figure 3. Predicted lifetime EOC risk for a female who has a PV in one of the high- or intermediate-risk genes included in the model, based on the different predictors of risk (risk factors (RFs) and PRS), for two family histories. Figures (a) and (b) show the lifetime risk for a carrier of a PV in *BRCA1*, Figures (c) and (d) show the lifetime risk for a carrier of a PV in *BRCA2*, Figures (e) and (f) show the lifetime risk for a carrier of a PV in *RAD51D*, Figures (g) and (h) show the lifetime risk for a carrier of a PV in *RAD51C*, while Figure (i) and (j) show the lifetime risk for a carrier of a PV in *BRIP1*. Figures (a), (c), (e), (g) and (i) show risks for an

unknown family history, while Figures (b), (d), (f), (h) and (j) show risks for a female whose mother is diagnosed with EOC at age 50. Predictions based on UK ovarian cancer incidences.

Figure 4. Calibration of the absolute and relative predicted five-year EOC risks, showing the observed and expected risks by quintile. The bars show the 95% confidence intervals for the observed risks. Relative risks were calculated relative to the overall mean of observed and predicted risks.

Figure 1

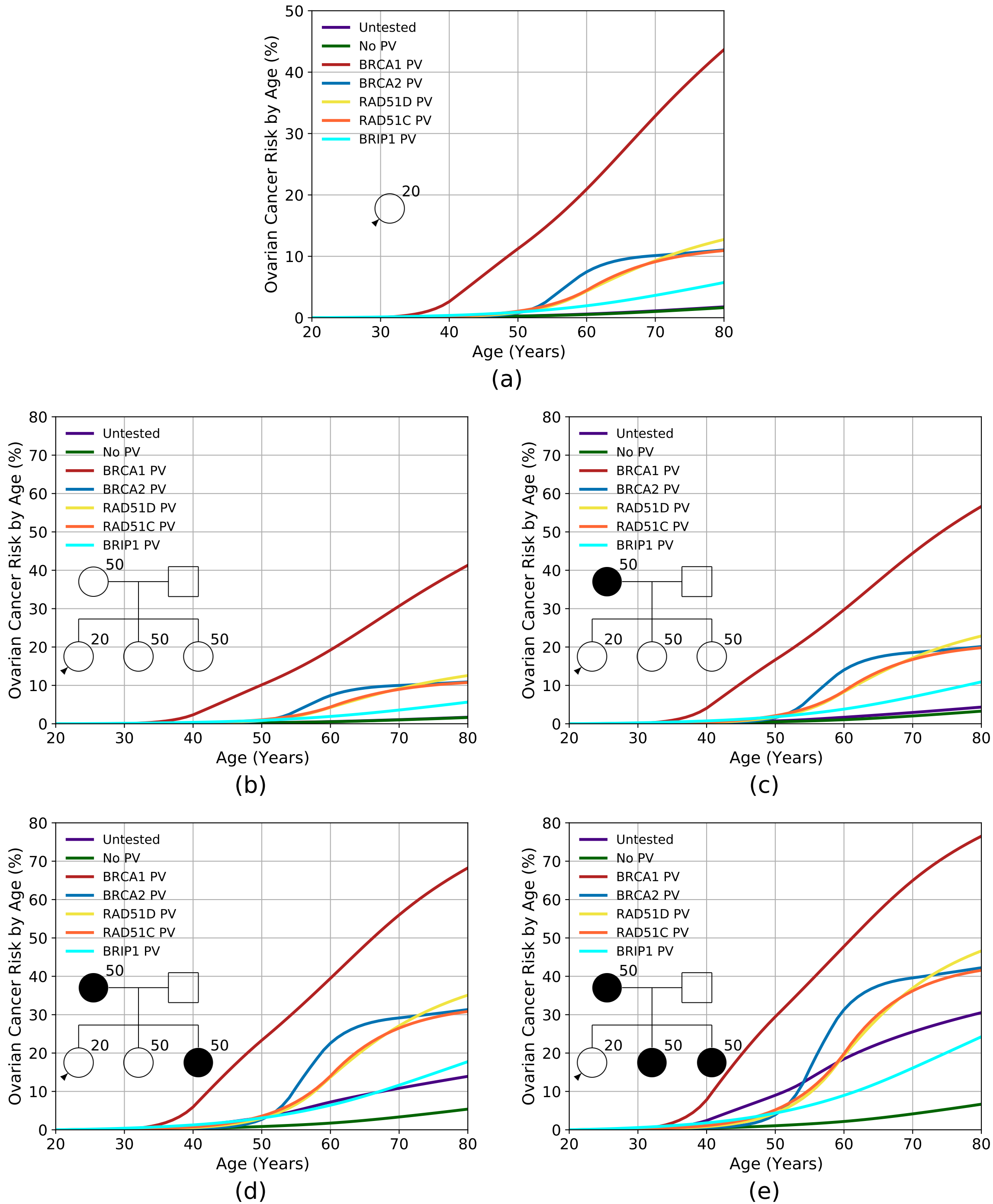

Untested
No PV
BRCA1 PV
BRCA2 PV
RAD51D PV
RAD51C PV
BRIP1 PV
Ovarian Cancer Risk by Age (%)
Age (Years)
20
50
(a)
(b)
(c)
(d)
(e)

Figure 2

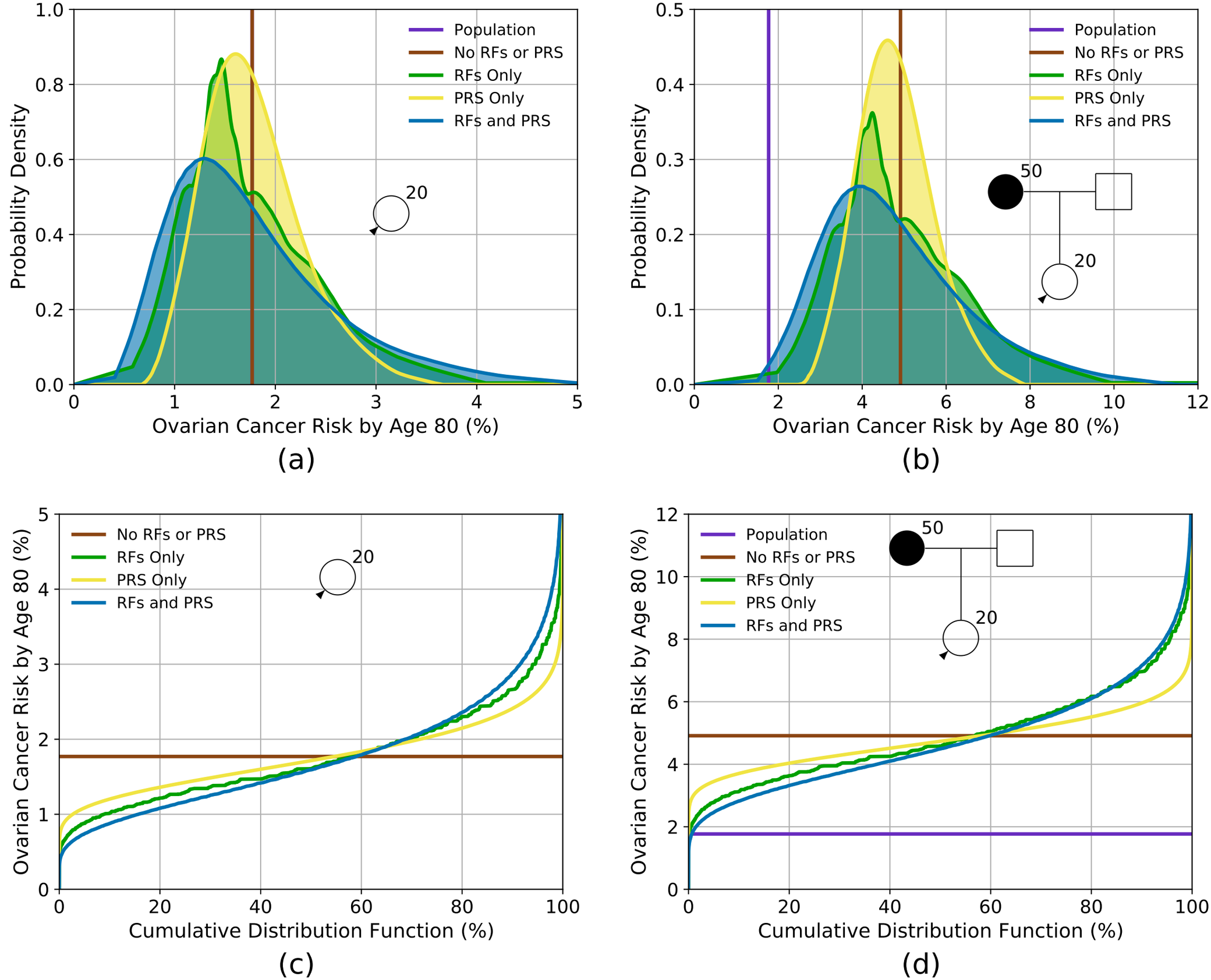

Population
No RFs or PRS
RFs Only
PRS Only
RFs and PRS
Probability Density
Ovarian Cancer Risk by Age 80 (%)
20
50
(a)
(b)
Ovarian Cancer Risk by Age 80 (%)
Cumulative Distribution Function (%)
(c)
(d)

Figure 3

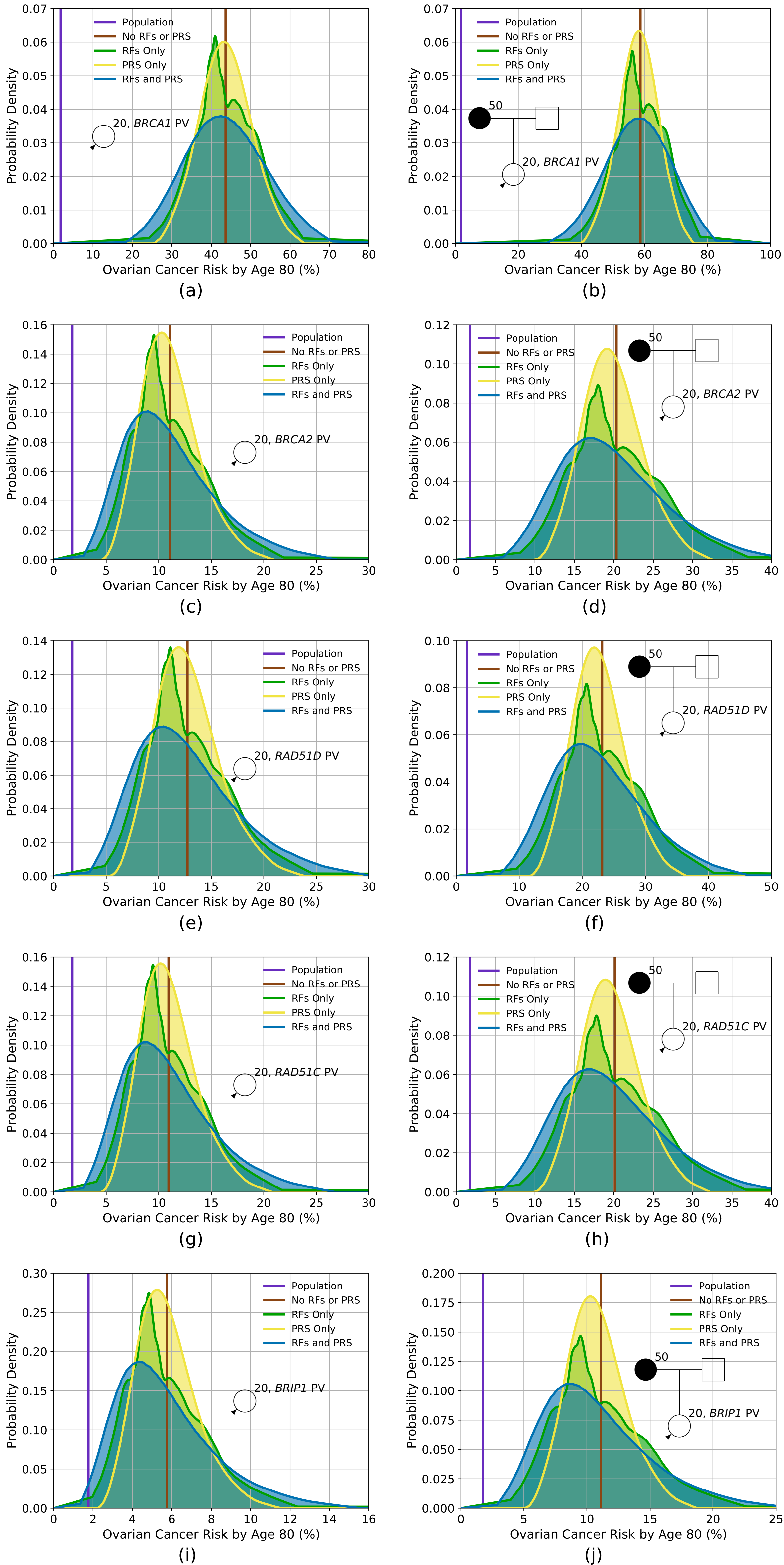
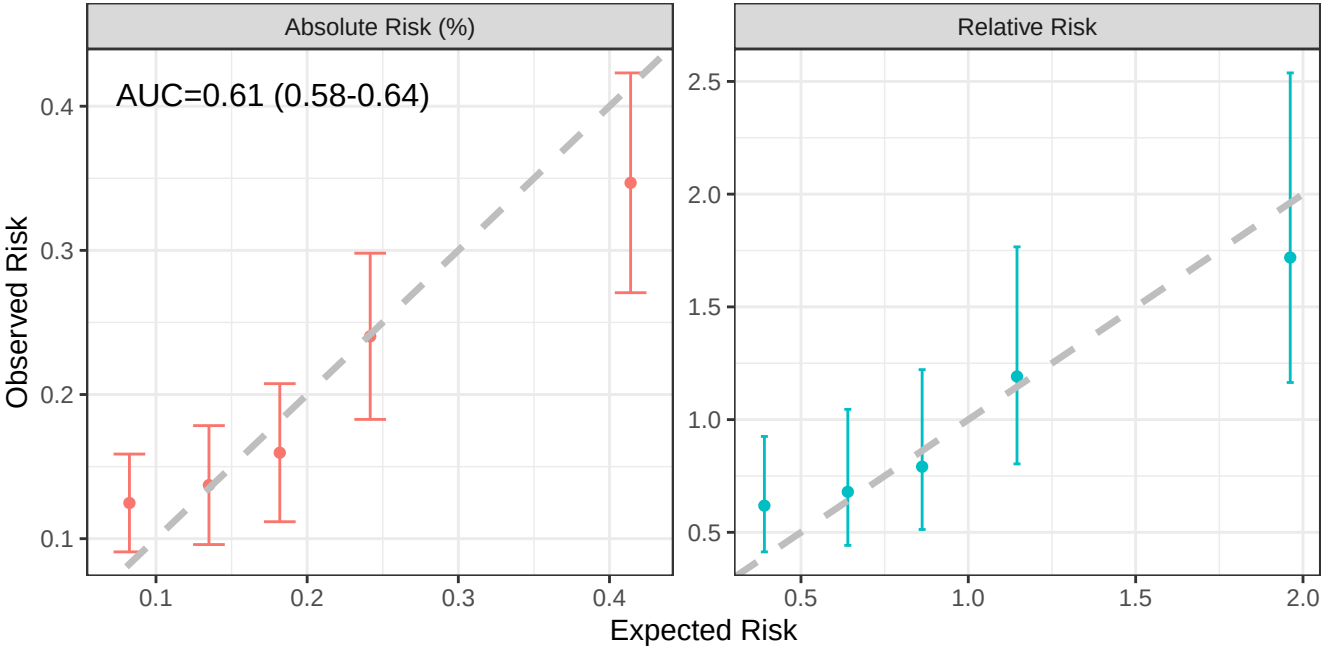

Population
No RFs or PRS
RFs Only
PRS Only
RFs and PRS
20, BRCA1 PV
50
20, BRCA2 PV
20, RAD51D PV
20, RAD51C PV
20, BRIP1 PV
Probability Density
Ovarian Cancer Risk by Age 80 (%)
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)

Figure 4


Absolute Risk (%)
Relative Risk
AUC=0.61 (0.58-0.64)
Observed Risk
Expected Risk
0.1
0.2
0.3
0.4
0.5
1.0
1.5
2.0
2.5